

# From Dust to Planetesimals: Criteria for Gravitational Instability of Small Particles in Gas

Ji-Ming Shi<sup>1,2</sup> and Eugene Chiang<sup>1,2,3</sup>  
jmshi@berkeley.edu

## ABSTRACT

Dust particles sediment toward the midplanes of protoplanetary disks, forming dust-rich sublayers encased in gas. What densities must the particle sublayer attain before it can fragment by self-gravity? We describe various candidate threshold densities. One of these is the Roche density, which is that required for a strengthless satellite to resist tidal disruption by its primary. Another is the Toomre density, which is that required for de-stabilizing self-gravity to defeat the stabilizing influences of pressure and rotation. We show that for sublayers containing aerodynamically well-coupled dust, the Toomre density exceeds the Roche density by many (up to about 4) orders of magnitude. We present 3D shearing box simulations of self-gravitating, stratified, dust-gas mixtures to test which of the candidate thresholds is relevant for collapse. All our simulations indicate that the larger Toomre density is required for collapse. This result is sensible because sublayers are readily stabilized by pressure. Sound-crossing times for thin layers are easily shorter than free-fall times, and the effective sound speed in dust-gas suspensions decreases only weakly with the dust-to-gas ratio (as the inverse square root). Our findings assume that particles are small enough that their stopping times in gas are shorter than all other timescales. Relaxing this assumption may lower the threshold for gravitational collapse back down to the Roche criterion. In particular, if the particle stopping time becomes longer than the sound-crossing time, sublayers may lose pressure support and become gravitationally unstable.

*Subject headings:* hydrodynamics — instabilities — planets and satellites: formation — protoplanetary disks — methods: numerical

## 1. INTRODUCTION

Gravitational instability is an attractive mechanism to form planetesimals, but how it is triggered in protoplanetary disks remains unclear. In one proposed sequence of events, most of the disk’s solids first coagulate into particles 0.1–1 m in size at orbital distances of a few AU. These “boulder”-sized bodies then further concentrate by the aerodynamic streaming instability (Youdin & Goodman 2005; Johansen et al. 2007; Bai & Stone

2010; and references therein). Local densities are so strongly enhanced by the streaming instability that they can exceed the Roche density (see §1.3 for a definition), whereupon collections of boulders may undergo gravitational collapse into more massive, bound structures.

A weakness of this scenario is that it presumes that particle-particle sticking (i.e., chemical adhesion) can convert most of the disk’s solids into boulders, or more accurately, particles whose momentum stopping times in gas

$$t_{\text{stop}} \equiv \frac{mv_{\text{rel}}}{F_{\text{drag}}} \quad (1)$$

are within a factor of 10 of the local dynamical time  $\Omega^{-1}$ , where  $\Omega$  is the Kepler orbital frequency,  $m$  is the particle mass,  $v_{\text{rel}}$  is the relative gas-particle velocity, and  $F_{\text{drag}}$  is the drag force whose

<sup>1</sup>Department of Astronomy, UC Berkeley, Hearst Field Annex B-20, Berkeley, CA 94720-3411

<sup>2</sup>Center for Integrative Planetary Science, UC Berkeley, Hearst Field Annex B-20, Berkeley, CA 94720-3411

<sup>3</sup>Department of Earth and Planetary Science, UC Berkeley, 307 McCone Hall, Berkeley, CA 94720-4767

form varies with disk environment (see, e.g., [Weidenschilling 1977](#)). Figure 1 relates  $t_{\text{stop}}$  to particle radius  $s$  as a function of disk radius  $r$  in the minimum-mass solar nebula (MMSN). For  $r \approx 1$ –10 AU, the condition  $\Omega t_{\text{stop}} = 0.1$ –1 corresponds to  $s \approx 0.1$ –1 m.

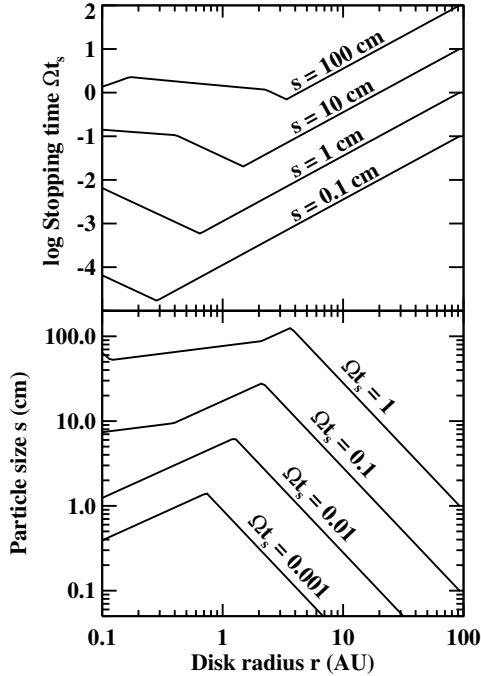


Fig. 1.— Stopping times of particles in the MMSN, normalized to the local dynamical time  $\Omega^{-1}$ . Disk parameters are taken from [Chiang & Youdin \(2010\)](#). Particles are assumed spherical with bulk density  $1 \text{ g/cm}^3$ . The kinks in the curves are due to transitions between different drag force laws as taken from [Weidenschilling \(1977\)](#); note that the transition between the Stokes and Epstein drag laws occurs when the gas mean free path equals  $2/9$  of the particle radius, not  $4/9$  as misprinted in that article). Marginally coupled particles ( $\Omega t_{\text{stop}} \sim 1$ ) correspond to meter-sized boulders at  $r \sim 1$  AU; decimeter-sized rocks at  $r \sim 10$  AU; and cm-sized pebbles at  $r \sim 100$  AU. The top panel plots  $\Omega t_{\text{stop}}$  at various fixed particle radii  $s$ ; the bottom panel plots the same data but at fixed  $\Omega t_{\text{stop}}$ . In this paper we are interested in the small particle  $\Omega t_{\text{stop}} \ll 1$  limit.

Unfortunately, particle-particle sticking might not produce boulders in sufficient numbers for the streaming instability to be significant. A

comprehensive study by [Zsom et al. \(2011\)](#); see also [Birnstiel et al. 2010](#)) found that for realistic, experiment-based sticking models that include both bouncing and fragmentation, particles no larger than  $\sim 1$  cm can form by sticking — even when the disk is assumed to have zero turbulence. According to Table 1 of [Zsom et al. \(2011\)](#), coagulation models over most of parameter space produce  $\tau_s \sim 10^{-4}$ – $10^{-2}$ . This range is too small for the streaming instability to concentrate particles strongly—see [Bai & Stone \(2010\)](#), who showed that when half or more of the disk’s solid mass has  $\Omega t_{\text{stop}} < 0.1$ , densities enhanced by the streaming instability still fall short of the Roche density by more than a factor of 10. Even if particle-particle sticking could grow bodies with  $\Omega t_{\text{stop}} \sim 0.1$ –1 (e.g., [Okuzumi et al. 2009](#), who neglected fragmentation), the disk’s solids may not be transformed into such bodies all at once. Rather, boulders may initially comprise a minority on the extreme tail of the size distribution. Unless they can multiply from a minority to a majority within the time it takes for them to drift radially inward by gas drag ( $\sim 100$ – $1000$  yr starting at 1 AU; [Weidenschilling 1977](#)), they threaten to be lost from the nebula by drag.

We are therefore motivated to ask whether gravitational instability is practicable for particles having realistically small sizes and concomitantly short stopping times, say  $\Omega t_{\text{stop}} \lesssim 10^{-2}$ . Smaller particles suffer the disadvantage that they are harder to concentrate; since they are well-entrained in gas, turbulence in the gas can loft particles above the midplane and prevent them from collecting into regions of higher density. The streaming instability provides one source of turbulence. Another driver of turbulence is the Kelvin-Helmholtz instability, caused by vertical velocity gradients which steepen as dust settles into a thin, dense “sublayer” at the disk midplane ([Weidenschilling 1980](#)). Several recent studies ([Chiang 2008](#), [Lee et al. 2010a,b](#); see also [Weidenschilling 2006, 2010](#)) have measured the maximum sublayer densities permitted by the Kelvin-Helmholtz instability. Neglecting self-gravity, they found that dust-to-gas ratios between  $\sim 2$ – $30$  are possible in disks that are locally enriched in metallicity by factors of  $1$ – $4$  above solar. Such local enrichment can be generated by radial drifts of particles relative to gas (see [Chiang & Youdin 2010](#) for a review).

For observational evidence of radial segregation of dust from gas, see [Andrews et al. \(2012\)](#).

Are such enhancements in the local dust-to-gas ratio sufficient to spawn planetesimals? How high must dust + gas densities be before the effects of self-gravity manifest? Our paper addresses these questions in the limit  $\Omega t_{\text{stop}} \ll 1$ , i.e., in the limit that particles are small enough to be well coupled to gas. In the next two subsections, we derive critical densities for gravitational instability in the cases of a pure gas disk (§1.1), and a disk composed of both gas and perfectly entrained ( $\Omega t_{\text{stop}} \rightarrow 0$ ) dust (§1.2). The two cases give remarkably different answers for dust-rich sublayers. In §1.3 we add two more densities from the literature to the list of proposed criteria for gravitational collapse. Table 1 summarizes the various candidate threshold densities.

In §2–§3, we present numerical simulations of 3D, self-gravitating, compressible flows of thin, dense sublayers of dust. We use these simulations to try to identify which of the proposed criteria (if any) is the most relevant for gravitational instability. Section 4 summarizes our findings but also points out the limitations of our numerical simulations, which are restricted to the asymptotic limit  $\Omega t_{\text{stop}} \rightarrow 0$ . We argue in §4.1 how finite but still small values of  $\Omega t_{\text{stop}}$  may lower the threshold for gravitational collapse.

### 1.1. Critical Density for Gravitational Instability in a Pure Gas Disk

The usual criterion for gravitational instability in a razor-thin pure gas disk is expressed in terms of the dimensionless parameter

$$Q_g \equiv \frac{c_g \Omega}{\pi G \Sigma_g} \quad (2)$$

where  $G$  is the gravitational constant,  $c_g$  is the gas sound speed, and  $\Sigma_g$  is the gas surface density ([Goldreich & Lynden-Bell 1965](#); [Toomre 1964](#); [Toomre 1981](#)). In (2), the Kepler orbital frequency  $\Omega$  has been substituted for the radial epicyclic frequency. If

$$Q_g < Q_g^* = 1, \quad (3)$$

the disk is gravitationally unstable to axisymmetric perturbations in the disk plane. The  $Q$ -criterion is a measure of the competition between stabilizing pressure, stabilizing rotation,

and de-stabilizing self-gravity (see, e.g., [Binney & Tremaine 2008](#)). When  $Q_g > 1$ , horizontal perturbations having lengthscales  $< 2c_g/G\Sigma_g$  are stabilized by pressure, while those having lengthscales  $> 2c_g/G\Sigma_g$  are stabilized by rotation. When  $Q_g$  equals 1, the first axisymmetric mode to become unstable to self-gravity has radial wavelength  $2c_g/G\Sigma_g$ . And as  $Q_g$  approaches 1 from above, the disk is increasingly susceptible to nonaxisymmetric perturbations which swing amplify ([Goldreich & Lynden-Bell 1965](#)).

The criterion  $Q_g \lesssim Q_g^*$  for gravitational instability can be translated into a criterion for the midplane density  $\rho_{g0}$  (the subscript “0” denotes the initial midplane value). We define a disk half-thickness  $H_g$  using

$$\Sigma_g \equiv 2\rho_{g0}H_g. \quad (4)$$

We also define a half-thickness  $H_g^\dagger$  using the usual relation from vertical hydrostatic equilibrium:

$$H_g^\dagger \equiv c_g/\Omega. \quad (5)$$

Ordinarily  $H_g \approx H_g^\dagger$  and we would not bother to distinguish the two; however, we will find cases later on where they differ by factors of several because of the effects of dust, and thus we take care to separate the two lengths now. Upon substitution of (4) and (5), the relation  $Q_g \lesssim Q_g^*$  is shown to be equivalent to

$$\rho_{g0} \gtrsim \rho_I^* = \frac{1}{2\pi} \frac{1}{Q_g^*} \frac{H_g^\dagger}{H_g} \rho^\dagger \quad (6)$$

where we have defined a reference density<sup>1</sup>

$$\rho^\dagger \equiv M_*/r^3 \quad (7)$$

with  $M_*$  and  $r$  equal to the mass of the central star and the disk radius, respectively.

The  $\rho_I^*$ -criterion (6) is sometimes used (e.g., [Lee et al. 2010a,b](#)) to signal gravitational instability in dusty gas disks (with  $\rho_{g0}$  replaced by the total dust + gas density  $\rho_{d0} + \rho_{g0}$ ,  $Q_g^* = 1$ , and  $H_g^\dagger/H_g = 1$ ). But using  $\rho_I^*$  for dust-gas mixtures is suspect because the criterion does not account explicitly for the two-phase nature of such media. In the next subsection we make such an accounting to derive a substantially different criterion for gravitational collapse.

<sup>1</sup>In this paper, we will superscript critical threshold densities with \*, and fiducial or reference quantities with †.

TABLE 1  
CANDIDATE CRITICAL DENSITIES FOR GRAVITATIONAL COLLAPSE

Critical Density	Value	Comment	Reference
$\rho_I^*$	$\frac{1}{2\pi} \frac{1}{Q_g^*} \frac{H_g^\dagger}{H_g} \frac{M_*}{r^3} \sim 0.16 \frac{M_*}{r^3}$ <sup>(a)</sup>	Equivalent to $Q_g < Q_g^*$ for pure gas disks	This paper, equation (6)
$\rho_{\text{Sekiya}}^*$	$0.60 \frac{M_*}{r^3}$	Required for the onset of an incompressible, axisymmetric overstable mode	<a href="#">Sekiya (1983)</a>
$\rho_{\text{Roche}}^*$	$3.5 \frac{M_*}{r^3}$	Required by satellite to resist tidal disruption by primary	<a href="#">Chandrasekhar (1987)</a>
$\rho_{\text{II}}^*$	$\frac{1}{2\pi} \frac{Q_g}{Q_d^{*2}} \left( \frac{\Sigma_g}{\Sigma_d} \right)^2 \frac{H_g^\dagger}{H_g} \frac{M_*}{r^3} \sim 2 \times 10^4 \frac{M_*}{r^3}$ <sup>(b)</sup>	Equivalent to $Q_d < Q_d^*$ for dust-rich sublayers in gas	This paper, equation (13)

<sup>(a)</sup> Value is derived for  $Q_g^* = 1$  and  $H_g^\dagger/H_g = 1$ .

<sup>(b)</sup> Value is derived for  $Q_d^* = 1$ ,  $Q_g = 30$ ,  $\Sigma_d/\Sigma_g = 0.015$ , and  $H_g^\dagger/H_g = 1$ .

## 1.2. Critical Density for Gravitational Instability in a Dust-Rich Sublayer in the Limit $\Omega t_{\text{stop}} \rightarrow 0$

For disks of gas and dust, gravitational instability should still be determined by the  $Q$ -criterion, except there is now the possibility that disk self-gravity is dominated by dust in a vertically thin sublayer at the midplane:

$$Q_d \equiv \frac{c_d \Omega}{\pi G \Sigma_d} \lesssim Q_d^* \text{ for instability.} \quad (8)$$

In using the dust surface density  $\Sigma_d$  in (8), we neglect the contribution of gas to the total surface density of the sublayer. Under typical circumstances, the error accrued is small.

In the limit  $\Omega t_{\text{stop}} \rightarrow 0$ , the dust-gas mixture represents a colloidal suspension. In this suspension, dust does not contribute to the pressure  $P$  — which is still provided entirely by gas — but instead adds to the inertia. In other words,

$$P = \rho_g c_g^2 = (\rho_g + \rho_d) c_d^2 \quad (9)$$

by definition of  $c_d$ , the speed of sound in the suspension:

$$c_d = \frac{c_g}{\sqrt{1 + \mu}} \quad (10)$$

where  $\rho_g$  is the local gas density,  $\rho_d$  is the local dust density, and  $\mu = \rho_d/\rho_g$  is the dust-to-gas ratio. In effect, dust increases the mean molecular weight of the gas.

Inserting (10) into (8) and using

$$\rho_{g0} = \rho^\dagger \frac{1}{2\pi Q_g} \frac{H_g^\dagger}{H_g}, \quad (11)$$

we solve for the total midplane density required for gravitational instability:

$$\rho_0 = \rho_{d0} + \rho_{g0} \gtrsim \rho_{\text{II}}^* \quad (12)$$

where

$$\rho_{\text{II}}^* = \frac{1}{2\pi} \frac{Q_g}{Q_d^{*2}} \left( \frac{\Sigma_g}{\Sigma_d} \right)^2 \frac{H_g^\dagger}{H_g} \rho^\dagger \quad (13)$$

$$\approx 2 \times 10^4 \rho^\dagger \left( \frac{Q_g}{30} \right) \left( \frac{1}{Q_d^*} \right)^2 \left( \frac{0.015}{\Sigma_d/\Sigma_g} \right)^2 \left( \frac{H_g^\dagger/H_g}{1} \right). \quad (14)$$

In (14), our normalizations for  $Q_g$  and the bulk (height-integrated but local to  $r$ ) metallicity  $\Sigma_d/\Sigma_g$  derive from the MMSN at  $r = 1$  AU ([Chiang & Youdin 2010](#)). For these parameter choices, the critical midplane density  $\rho_{\text{II}}^*$  is an astonishing five orders of magnitude greater than  $\rho_I^*$ . It is possible that real disks have masses and bulk metallicities enhanced over the MMSN by factors of a few, in which case  $\rho_{\text{II}}^*$  would be larger than  $\rho_I^*$  by about three orders of magnitude.

### 1.3. Other Critical Densities

Another threshold density, already alluded to at the beginning of §1, is the Roche density:

$$\rho_{\text{Roche}}^* = 3.5 \frac{M_*}{r^3}. \quad (15)$$

The Roche density is the density required for a strengthless, incompressible, fluid body in hydrostatic equilibrium to resist tidal disruption, when in synchronous orbit at distance  $r$  about a star of mass  $M_*$  (e.g., [Chandrasekhar 1987](#)).

Yet another candidate threshold was proposed by [Sekiya \(1983\)](#), who found that when the mid-plane density exceeds

$$\rho_{\text{Sekiya}}^* = 0.60 \frac{M_*}{r^3}, \quad (16)$$

the disk becomes susceptible to an unstable, incompressible, axisymmetric mode in which in-plane motions generate out-of-plane bulges (i.e., an annulus that contracts radially becomes thicker vertically, and vice versa). The nonlinear outcome of this instability is not known, but [Sekiya \(1983\)](#) speculated that the dust sublayer might eventually fragment on the scale of the wavelength of the overstable mode, and that dust particles might sediment toward the centers of fragments to form the first-generation planetesimals.

Table 1 summarizes the four candidate threshold densities. For realistic parameters ( $Q_g \sim 10$ – $30$ ;  $\Sigma_d/\Sigma_g \sim 0.015$ – $0.15$ ), the four densities obey

$$\rho_I^* < \rho_{\text{Sekiya}}^* < \rho_{\text{Roche}}^* \ll \rho_{\text{II}}^*. \quad (17)$$

The smallest three densities in this hierarchy are fixed multiples of the reference density  $\rho^\dagger = M_*/r^3$  (with coefficients  $\sim 1/2\pi \approx 0.16$ , 0.6, and 3.5, respectively). The last density  $\rho_{\text{II}}^*$  can, in principle, be arbitrarily larger than  $\rho^\dagger$ ; for typical, astrophysically plausible parameters, it is 2–4 orders of magnitude larger.

Which of the four densities in Table 1 is the most accurate predictor of gravitational collapse? In the next two sections, we describe numerical simulations performed in the  $\Omega t_{\text{stop}} \rightarrow 0$  limit that attempt to answer this question. We will find unfortunately that the numerical expense of simulating thin sublayers of dusty gas will force us into a parameter space where the difference between  $\rho_{\text{II}}^*$  and the other densities is not as large as it is in reality; we will have to make do with what we can.

## 2. METHODS

### 2.1. Code

We simulate hydrodynamic, self-gravitating, stratified flows in disks using **Athena**, configured for a shearing box, with no magnetic fields ([Stone et al. 2008](#); [Stone & Gardiner 2010](#)). Dust is assumed to be perfectly aerodynamically coupled to gas so that they share the same velocity field  $\mathbf{v}$ .

The equations solved are:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (18)$$

$$\frac{\partial \rho_g}{\partial t} + \nabla \cdot (\rho_g \mathbf{v}) = 0, \quad (19)$$

$$\begin{aligned} \frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v} \mathbf{v} + \mathbf{P}) &= -\rho \nabla \Phi \\ -2\rho(\Omega \hat{\mathbf{z}}) \times \mathbf{v} + 2q\rho\Omega^2 x \hat{\mathbf{x}} - \rho\Omega^2 z \hat{\mathbf{z}}, \end{aligned} \quad (20)$$

$$\nabla^2 \Phi = 4\pi G \rho, \quad (21)$$

where  $\rho = \rho_g + \rho_d$  is the total density of the dust-gas suspension,  $\mathbf{P} = P\mathbf{I}$  is a diagonal tensor with components  $P = \rho_g c_g^2$  as defined in equation (9) with constant  $c_g$  (isothermal approximation),  $\Omega$  is the mean (constant) orbital frequency,  $\hat{\mathbf{x}}$  points in the radial direction,  $\hat{\mathbf{z}}$  points in the vertical direction, and  $\Phi$  is the self-gravitational potential of the dust-gas mixture. We choose the shear parameter  $q = 3/2$  for Keplerian flow.

#### 2.1.1. Algorithms and boundary conditions

**Athena 4.0** provides several schemes for time integration and spatial reconstruction, and for solving the Riemann problem. Having experimented with various options, we adopted the van Leer algorithm for our dimensionally unsplit integrator ([van Leer 2006](#); [Stone & Gardiner 2009](#)); a piecewise linear spatial reconstruction in the primitive variables; and the HLLC (Harten-Lax-van Leer-Contact) Riemann solver. To account for disk self-gravity, we use the routines written by [Koyama & Ostriker \(2009\)](#) and [Kim et al. \(2011\)](#) which solve Poisson’s equation using fast Fourier transforms.

Boundary conditions for our hydrodynamic flow variables (including density and velocity, but not the self-gravitational potential) are shearing-periodic in radius ( $x$ ) and periodic in azimuth ( $y$ ). For vertical height ( $z$ ), we experimented with both periodic and outflow boundary conditions,



and chose periodic boundary conditions to ensure strict mass conservation. When outflow boundary conditions were employed, mass was lost from the boundaries at early times and complicated the interpretation of our results. We verified that our results are insensitive to box height for sufficiently tall boxes; see §3 for explicit tests.

The Poisson solver implements shear-periodic boundary conditions in  $x$ , periodic boundary conditions in  $y$ , and vacuum boundary conditions in  $z$  (Koyama & Ostriker 2009; Kim et al. 2011). In our simulations, self-gravity is dominated by dust, and our boxes are tall enough to contain the entire dust layer. Both vertical and radial stellar tidal gravity are included as source terms in the van Leer integrator.

We further augmented the code to include dust in the limit of zero stopping time. In this limit, dust shares the same velocity field as gas, and contributes only to the mass density. In our modified version of **Athena**, two continuity equations are solved: one for the entire mixture ( $\rho = \rho_g + \rho_d$ , see equation 18), and one for the gas ( $\rho_g$ , see equation 19). The dust density is given by the difference ( $\rho - \rho_g$ ). The remaining hydrodynamic equations govern the dust-gas mixture ( $\rho$ ), but with gas ( $\rho_g$ ) contributing solely to the pressure  $P$  (see Equation (20)). For simplicity, we adopt an isothermal equation of state so that  $P$  is related to  $\rho_g$  by equation (9) for constant  $c_g$ . Isothermal flows are more prone to gravitational instability than adiabatic ones (Mamatsashvili & Rice 2010).

We also modified the HLLC Riemann solver to accommodate our dust-gas mixture. Changes include the following: (1) The speeds of the left, right, and contact waves are reduced by a factor  $(1 + \mu)^{-1/2}$ , where  $\mu \equiv \rho_d/\rho_g$  is the local dust-to-gas ratio, to account for the added inertia from dust (see equation 10). (2) The pressure in the contact region is replaced by an equivalent but numerically more accurate form based on equation (10.76) in Toro (1999). (3) When calculating left/right momentum fluxes, we ensure that only gas contributes to the pressure by using  $\rho_g$  and not  $\rho$ . (4) For the flux solver to predict the pressure and wave speeds, the left/right gas densities require specification. We therefore add a reconstruction process for the gas density which interpolates cell-centered values to cell boundaries to second-order accuracy.

Previous studies of dust in the perfectly coupled limit (Chiang 2008; Lee et al. 2010a,b) also introduced a static background radial pressure gradient to mimic sub-Keplerian rotation of gas in a pressure-supported disk. We could also add the appropriate source term to the van Leer integrator. However, since our goal is to determine the minimum densities required for gravitational collapse and not to study vertical shear instabilities (i.e., the Kelvin-Helmholtz instability), we omit the background pressure gradient for simplicity.

In many of our simulations, the dust layer at the midplane collapses vertically because it is gravitationally unstable. Because of our boundary conditions, “fresh” gas from outside the simulation box cannot enter into the box, and thus in the event of gravitational collapse toward the midplane, the topmost and bottommost regions of our simulation domain become evacuated. Low-density gas in those regions become increasingly easy to accelerate, and the code timestep shortens by orders of magnitude, effectively halting the simulation. The dramatic reduction in timestep is not a serious limitation, as it usually occurs after the collapsing dust has attained some saturated state (see §3.2.1). In any case, we are more interested in the onset of gravitational instability than its nonlinear development.

### 2.1.2. Code tests

The following test problems helped to validate our code.

*Linear wave propagation.*— We propagated a small-amplitude 1D wave in a medium with a uniform background dust-to-gas ratio, with periodic boundary conditions, no background shear, and no gravity. We found the simulated wave speed matched the reduced sound speed calculated in (10). We chose our box to be one wavelength long, so that after one wave period, the wave crossed the boundaries and returned to its original position. With  $N = 128$  grid cells and an initial (fractional) wave amplitude  $A = 10^{-4}$ , we found the deviation  $\delta q \equiv \frac{1}{N} \sum_{i=1}^N |q_i - q_i^0| \approx 2 \times 10^{-8}$ , where  $q \in \{\rho_d, \rho_g, \rho\}$  and  $q^0$  represents the initial condition.

*Dust cloud advection.*— We advected a Gaussian-shaped dust cloud in a 1D domain. The cloud oc-

cupied about half the size of the box and the code was run for one box-crossing time. With  $N = 256$  grid cells, the root-mean-squared deviation in the shape of the cloud was  $< 1\%$ .

*Hydrostatic equilibrium of a stratified but non-self-gravitating dusty disk.*— Omitting self-gravity but including stellar gravity (both radial and vertical), we set up 3D dust-gas mixtures in hydrostatic equilibrium. A variety of vertical profiles for the dust-to-gas ratio were tested, ranging from uniform to linear to more complicated functional forms. All equilibria were found to be stable against small perturbations, even for dust-to-gas ratios as large as several hundred.

*Gravitational instability of 3D pure gas disks.*— We simulated isothermal, gravitationally unstable disks of pure gas in 3D. The gas was initialized in hydrostatic equilibrium (computed with vertical self-gravity), and box heights spanned approximately  $\pm 4$  initial gas scale heights. We found that  $Q_g = 1$  did not trigger gravitational instability, whereas  $Q_g = 0.5$  did. Our results are consistent with those of Goldreich & Lynden-Bell (1965), who found analytically that  $Q_g^* = 0.676$  for a finite-thickness isothermal gas disk.

## 2.2. Initial Conditions and Run Parameters

Initial conditions for our science simulations are of a dust-gas mixture with a pre-defined vertical profile for the dust-to-gas ratio  $\mu(z) = \rho_d(z)/\rho_g(z)$ . We choose the form

$$\mu(z) \equiv \mu_0 \operatorname{sech}^2 \left( \frac{z}{z_d} \right), \quad (22)$$

where  $\mu_0$  is the midplane dust-to-gas ratio. The scale height  $z_d$  can be thought of as the half-thickness of the dust layer insofar as  $\rho_g(z)$  is constant with  $z$ .

The isothermal dust-gas mixture is initialized in vertical hydrostatic equilibrium, including both stellar tidal gravity and disk self-gravity:

$$\frac{c_g^2}{\rho_g + \rho_d} \frac{d\rho_g}{dz} = -\Omega^2 z - 4\pi G \int_0^z (\rho_g + \rho_d) dz. \quad (23)$$

We solve numerically the differential form of (23).

Taking derivatives, we find

$$\frac{d}{dz} \left[ (1 + \mu)^{-1} \frac{d \ln \rho_g}{dz} \right] = -\frac{1}{H_g^\dagger} - \frac{4\pi G}{c_g^2} \rho_g (1 + \mu), \quad (24)$$

where  $H_g^\dagger \equiv c_g/\Omega$  is a fiducial (constant) gas scale height, not to be confused with any actual disk scale height. A non-dimensional form of (24) is given by

$$\frac{d}{d\tilde{z}} \left[ (1 + \mu)^{-1} \frac{d \ln \tilde{\rho}_g}{d\tilde{z}} \right] = -1 - \frac{2}{h_g Q_g} \tilde{\rho}_g (1 + \mu), \quad (25)$$

where we have defined the dimensionless variables  $\tilde{z} \equiv z/H_g^\dagger$ ,  $\tilde{\rho}_g \equiv \rho_g/\rho_{g0}$  (where  $\rho_{g0}$  is the midplane gas density), and  $h_g \equiv H_g/H_g^\dagger$ , with

$$H_g \equiv \Sigma_g/(2\rho_{g0}). \quad (26)$$

Upon insertion of (22), equation (25) can be solved numerically for  $\tilde{\rho}_g(\tilde{z})$ . But the solution must satisfy the following two constraints:

$$h_g = \int_0^\infty \tilde{\rho}_g(\tilde{z}) d\tilde{z} \quad (27)$$

by definition of  $H_g$ , and

$$\frac{\Sigma_d}{\Sigma_g} = \frac{1}{h_g} \int_0^\infty \tilde{\rho}_g(\tilde{z}) \mu(\tilde{z}) d\tilde{z} \quad (28)$$

for a fixed height-integrated (i.e., bulk) metallicity  $\Sigma_d/\Sigma_g$ .

Our procedure is as follows. We freely specify  $Q_g$ ,  $\Sigma_d/\Sigma_g$ , and  $\mu_0$  as model input parameters. We then iteratively solve equations (25), (27) and (28) for the three unknowns  $\tilde{\rho}_g(\tilde{z})$ ,  $h_g$ , and  $z_d$ . First we guess  $z_d$  and  $h_g$ , and integrate (25) to obtain  $\tilde{\rho}_g(\tilde{z})$ . If  $\tilde{\rho}_g$  so calculated fails (27), then we revise  $h_g$  and re-integrate (25), repeating until (27) is satisfied. Next we check (28). If  $\tilde{\rho}_g(\tilde{z})$  and  $h_g$  fail (28), then we revise  $z_d$  and repeat the procedure from the beginning, re-integrating (25) to obtain  $\tilde{\rho}_g$ , re-establishing (27), and so on. Typically  $\sim 100$  iterations ( $\sim 10$  for  $z_d \times \sim 10$  for  $h_g$ ) are required before all constraints are satisfied to  $\sim 1\%$  accuracy in  $z_d$  and  $10^{-6}$  accuracy in  $h_g$ .

Table 3 lists the parameters of our models. Note that these parameters do not describe plausible protoplanetary gas disks; in particular, our model metallicities  $\Sigma_d/\Sigma_g$  are orders of magnitude above the solar value of  $\sim 0.015$ . Parameters are

instead chosen to yield disk flows that our code can adequately resolve while still testing equation (14). Unfortunately, more astrophysically realistic parameters correspond to dust sublayers that are too vertically thin for us to resolve numerically; the code timestep, set by the sound-crossing time across a grid cell, becomes prohibitively short as thinner dust layers are considered. This difficulty means that the difference between  $\rho_{\text{II}}^*$  and the other candidate threshold densities is much less than what it is in reality, and our ability to distinguish between the candidates degrades as a result.

Figure 2 plots the initial conditions for our standard model ( $S = \text{STD32}$ ), for which  $Q_g = 24$ ,  $\Sigma_d/\Sigma_g = 8$ , and  $\mu_0 = 35$ . For this specific case, we calculate that  $h_g = 0.20 c_g/\Omega$  and  $z_d = 0.083 c_g/\Omega$ . The top and bottom boundaries of our simulation box are indicated by dotted vertical lines; typically box heights span  $\pm 4z_d$  (see §3 for box height tests). The right-hand panel of Figure 2 compares gas density profiles computed with and without self-gravity, and with and without dust, and shows that both the weight and self-gravity of the embedded dust layer force gas into a similarly thin layer.

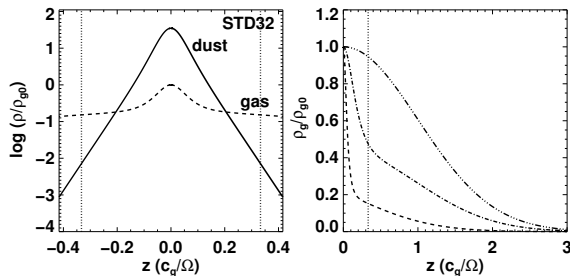


Fig. 2.— Left: Initial dust and gas densities for standard run  $S = \text{STD32}$ . Right: Initial gas density (dashed) on an expanded scale, together with the gas density computed without self-gravity but with dust (dash-dot) and without dust but with self-gravity (dash-double-dot). Vertical lines in both panels lines delimit the top and bottom of our computational box.

Figure 3 plots the initial force densities within the upper half of the simulation box to demonstrate how well vertical hydrostatic equilibrium is satisfied. The sum of stellar gravity (blue dashed curve) and disk self-gravity (red solid curve computed via the integral in equation 23) should equal

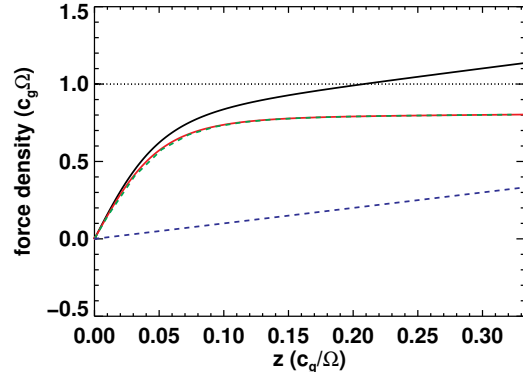


Fig. 3.— Vertical force balance for our initial conditions. The force density due to self-gravity is computed two ways: by direct integration of the density profile (red solid), and by using the code’s Poisson solver (green dashed). The two methods agree. The force density due to the pressure gradient (black solid) should equal the sum of self-gravity and the static stellar potential (blue dashed). The horizontal gray dotted line shows the ratio of pressure to gravity. All force densities are shown in their absolute values.

the pressure gradient (black solid curve). It does, as evidenced by the ratio of pressure to gravity (gray dotted line) which is practically constant at unity. We also overplot the self-gravitational force computed by our 3D Poisson solver (green dashed curve); the agreement with the exact solution is good.

Every simulation listed in Table 3 is perturbed from its initial equilibrium by adding random cell-to-cell fluctuations of amplitude  $\sim 10^{-3} c_g$  to the velocity field. The typical duration of a simulation is  $\sim 20 \Omega^{-1}$ . Our rationales for box size and resolution are explained in §3.

### 3. RESULTS

Results for 2D shearing sheets are described in §3.1, and those for 3D shearing boxes are in §3.2.

#### 3.1. 2D Shearing Sheet

For two-dimensional dusty disks, the criterion for gravitational instability reads

$$Q_d = \frac{c_d \Omega}{\pi G (\Sigma_d + \Sigma_g)} = \frac{Q_g}{(1 + \mu_0)^{3/2}} < Q_{d,2D}^* \quad (29)$$



TABLE 2  
2D SIMULATION PARAMETERS

Name	$Q_g$	$\mu_0$	$Q_d$	$\lambda_c(c_g/\Omega)$	$L_x \times L_y$ ( $c_g^2/\Omega^2$ )	Resolution	GI <sup>(a)</sup>	Duration ( $\Omega^{-1}$ )
S2D0	12	8	0.44	0.93	$10 \times 10$	$256 \times 256$	Y	5.9
S2D1	12	8	0.44	0.93	$10 \times 10$	$512 \times 512$	Y	6
S2D2	12	8	0.44	0.93	$0.5 \times 0.5$	$32 \times 32$	N	100
S2D3	12	8	0.44	0.93	$0.5 \times 0.5$	$128 \times 128$	N	100
S2D4	12	4.2	1.0	2.8	$30 \times 30$	$256 \times 256$	N	100
S2D5	12	2.3	2.0	6.9	$70 \times 70$	$256 \times 256$	N	100
S2D6	6	4.2	0.5	1.4	$15 \times 15$	$256 \times 256$	Y	6.2

<sup>(a)</sup> GI = Gravitational Instability. Y means  $\max \Sigma_d$  increases by orders of magnitude over a few dynamical times, and N means it does not.

We test this criterion by constructing a series of 2D shearing sheet simulations with various values of  $Q_g$  and  $\mu_0$ , thereby seeing if we can converge on a unique value for  $Q_{d,2D}^*$ . Although total surface densities can change during the simulation, the dust-to-gas ratio stays fixed at its initial value because of our perfect-coupling approximation. Initial conditions are as follows: for a given domain size  $L_x$  and  $L_y$ , the flow velocity  $\mathbf{v} = -\frac{3}{2}\Omega x \hat{e}_y$  and the surface density  $\Sigma = \Sigma_0 + \delta\Sigma \cos(\mathbf{k} \cdot \mathbf{x})$ , with  $\Sigma_0 = \Sigma_{g0} + \Sigma_{d0} = \Sigma_{g0}(1 + \mu_0)$ ,  $\delta\Sigma/\Sigma_0 = 0.01$ ,  $k_x = -2(2\pi/L_x)$ , and  $k_y = 2\pi/L_y$ . In our 2D simulations, we choose  $c_g = \Omega = \Sigma_{g0} = 1$  as our units.

Table 2 lists the parameters for our 2D runs. Our standard 2D run, labeled S2D0, has  $Q_g = 12$  and  $\mu_0 = 8.0$  and therefore  $Q_d = 0.44$ . For this run, the domain size is chosen large enough to easily fit the critical wavelength  $\lambda_c$  for gravitational instability:  $L_x = L_y = 10c_g/\Omega \gtrsim 10\lambda_c$ , where

$$\lambda_c \equiv \frac{2c_d^2}{G\Sigma_0} \quad (30)$$

is the wavelength of the fastest growing mode according to the WKB dispersion relation for axisymmetric waves. It is also the wavelength of the first mode to become unstable when  $Q_d$  just crosses  $Q_{d,2D}^*$ . The resolution of the standard run is  $N_x \times N_y = 256 \times 256$  so that one critical wavelength is resolved across  $\sim 10$  grid cells.

For S2D0, we find that the disk is indeed gravitationally unstable: density waves steepen quickly, and dense clumps of dusty gas form before one orbital period elapses. A simple way to portray instability is to track the maximum dust density

$\max \Sigma_d$  versus time — this is done in Figure 4, which shows that the maximum dust density increases by two orders of magnitude over a few dynamical times for our standard run.

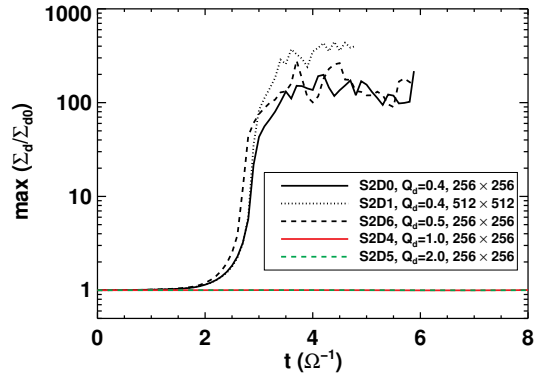


Fig. 4.— Time evolution of the maximum dust density  $\max \Sigma_d$  for our 2D shearing sheet simulations. The critical value  $Q_{d,2D}^*$  below which gravitational instability is triggered appears to be between 0.5 and 1.0.

Also shown in Figure 4 are results for other runs. In S2D4, S2D5, and S2D6, either  $Q_g$  or  $\mu_0$  is varied relative to our standard run, so that  $Q_d$  varies from 0.5 to 2.0. For all these runs, the domain size is  $\sim 10\lambda_c$  in each direction and the resolution is  $\sim 10$  cells per  $\lambda_c$ , just as in the standard case. Taken together, the results indicate that

$$0.5 < Q_{d,2D}^* < 1.0. \quad (31)$$

Other runs explore the effects of varying resolution and domain size. Doubling both  $N_x$  and  $N_y$

relative to our standard run (S2D1) enables higher maximum densities to be achieved when the instability saturates, but otherwise does not seem to alter the evolution. Reducing the size of the box so that it can no longer accommodate even a single critical wavelength (S2D2, S2D3) results in no instability, as expected (Gammie 2001; Johnson & Gammie 2003).

### 3.2. 3D Stratified Dusty Disks

Equation (8; equivalently 29) gives the criterion for gravitational instability in a 2D razor-thin sheet. For a 3D, vertically stratified disk, there is some ambiguity as to how we evaluate  $c_d$  in equation (8) because its value varies with height. Here we simply take  $c_d$  to be its value at the midplane, so that criterion (8) becomes

$$Q_d \simeq Q_g \frac{1}{\Sigma_d/\Sigma_g} \frac{1}{(1 + \mu_0)^{1/2}} \lesssim Q_d^*. \quad (32)$$

An alternative is to calculate a vertically averaged, density-weighted sound speed. We found, however, that such a procedure made little practical difference, since dust densities are much greater than gas densities near the midplane and drop steeply with height.

#### 3.2.1. Standard run ( $S = \text{STD32}$ )

To orient the reader, we present results for our standard 3D run (S, also labeled STD32 in §3.2.2), for which  $Q_d = 0.5$ . The full set of model S parameters are listed in Table 3, and the initial gas and dust density profiles are displayed in Figure 2. Our simulation box extends  $\pm 4z_d$  vertically, and  $14z_d$  in either horizontal direction. Each horizontal length is about twice the critical wavelength ( $\lambda_c \approx 6.3z_d$ ). The resolution is  $32 \times 32 \times 32$  so that one horizontal critical wavelength spans  $\sim 16$  cells, and one vertical scale length  $z_d$  spans 4 cells. These choices for domain size and resolution are tested in §3.2.2. The simulation is terminated at  $\sim 10.3\Omega^{-1}$ , at which point the timestep has become three orders of magnitude smaller than the initial timestep (see the final paragraph of §2.1.1).

Figure 5 displays a time series of the volume-rendered dust density in the bottom half of the box. Over the course of several dynamical times, density waves shear and amplify, eventually concentrating into a single azimuthally elongated filament. This filament then fragments radially. The

fragments gravitationally scatter and merge; by the end of the simulation, two clumps remain.

A simple diagnostic that we use throughout this paper is the time evolution of the maximum dust density, shown in the left panel of Figure 6. Comparison with Figure 5 reveals that  $\max \rho_d$  grows exponentially when the filament fragments radially. The maximum dust density ceases to rise once the clumps finish coalescing. At this point each clump is gravitationally bound, with a maximum central density that depends on the simulation resolution (§3.2.2).

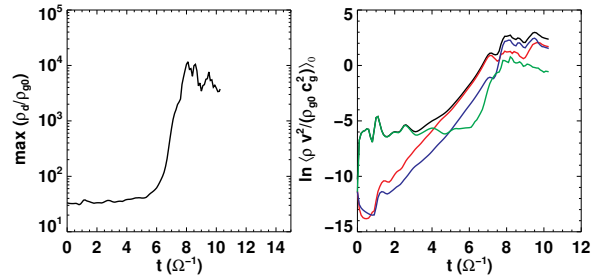


Fig. 6.— Left: Time evolution of maximum dust density for run STD32 ( $= S$ ). Right: Time evolution of kinetic energies averaged horizontally and vertically over a thin slab subtending two grid cells at the midplane (red =  $x$ -component of kinetic energy; blue =  $y$ ; green =  $z$ ; black = total).

The right panel of Figure 6 shows the time evolution of various kinetic energy densities, evaluated in the three directions and excluding the background Keplerian shear. The energy densities are averaged horizontally and vertically over a thin slab subtending two grid cells at the midplane (qualitatively similar results are obtained over larger vertical averages). The horizontal kinetic energies grow exponentially from  $t = 2 - 7\Omega^{-1}$ , with an exponential growth rate of  $\sim 1.5\Omega$ . Radial motions dominate azimuthal motions until the end of the simulation when they become comparable. Vertical motions develop immediately after the beginning of the simulation because our discretized initial conditions cannot be in perfect hydrostatic balance; however the magnitude of the vertical motions is small and stays roughly constant for  $t \lesssim 6\Omega^{-1}$ . For  $t \gtrsim 6\Omega^{-1}$ , vertical motions amplify but for the most part remain smaller than horizontal motions.

The in-plane motions of the dusty clumps are

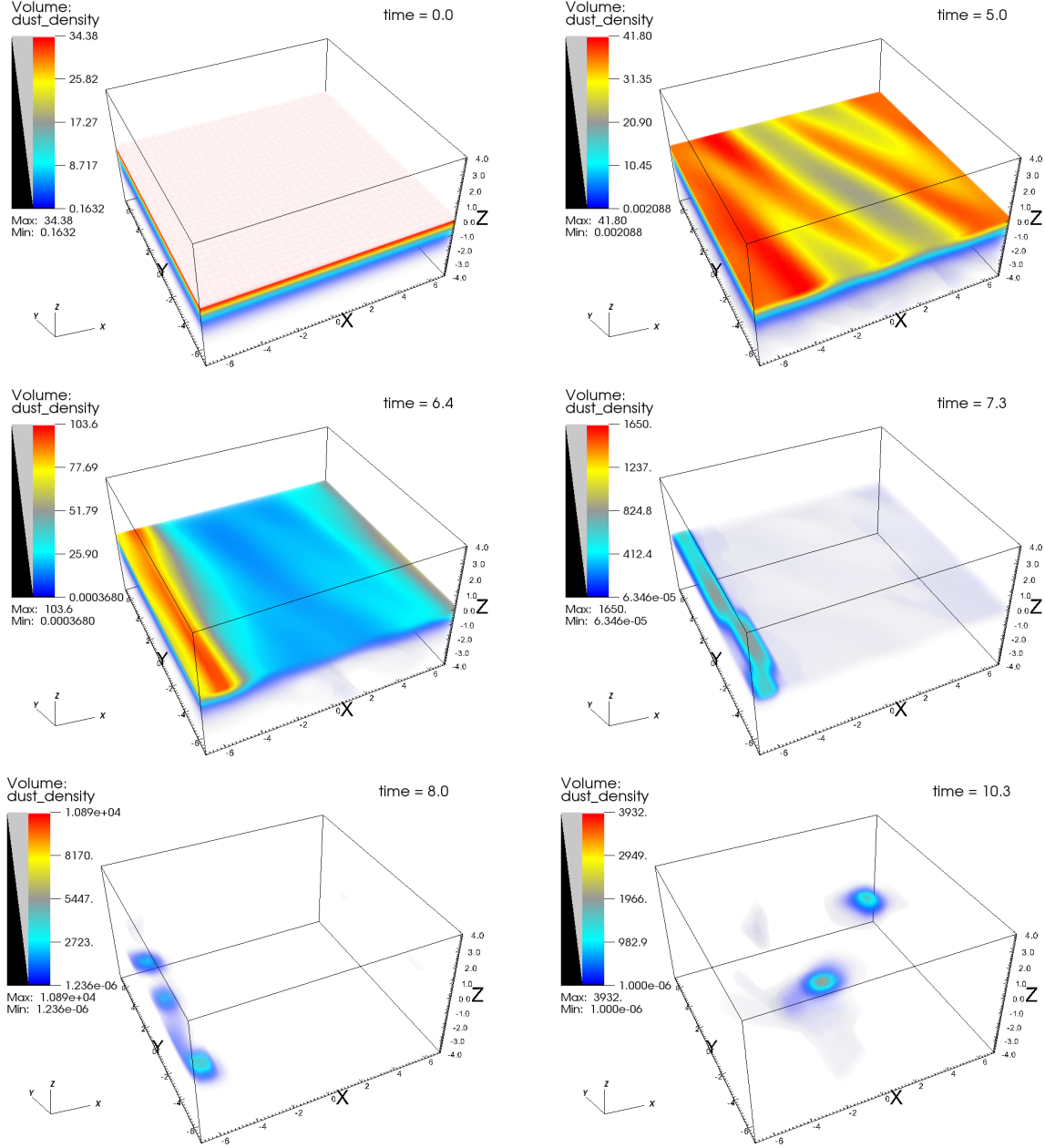


Fig. 5.— Time evolution of gravitational instability in our standard 3D stratified dusty disk (run S = STD32). Shown are volume renderings of dust density for the bottom half of the disk at  $t = 0, 5.0, 6.4, 7.3, 8.0,$  and  $10.3\Omega^{-1}$  (left to right, top to bottom).

TABLE 3  
3D SIMULATION PARAMETERS (“SCIENCE RUNS”)

Name	$Q_g$	$\mu_0$	$\Sigma_d/\Sigma_g$	$H_g$ ( $c_g/\Omega$ )	$z_d$ ( $c_g/\Omega$ )	$\lambda_c$ ( $z_d$ )	$L_x \times L_y \times L_z$ ( $z_d^3$ )	Resolution	Duration ( $\Omega^{-1}$ )	$Q_d$	$\rho_0$ ( $\rho^\dagger$ )	$\rho_I^{*(a)}$ ( $\rho^\dagger$ )	$\rho_{II}^{*(b)}$ ( $\rho^\dagger$ )	GI <sup>(c)</sup>
S	24	35.0	8.0	0.20	0.083	6.3	$14 \times 14 \times 8$	$32 \times 32 \times 32$	10.3	0.5	1.20	0.16	0.30	Y
R1	24	165.0	2.0	0.91	0.011	42.6	$90 \times 90 \times 8$	$256 \times 256 \times 32$	11	0.93	1.21	0.16	1.05	Y/N <sup>(d)</sup>
R2	12	93.0	0.67	1.04	0.008	150.3	$256 \times 256 \times 8$	$256 \times 256 \times 32$	30	1.86	1.20	0.16	4.09	N
R3	12	143.0	0.54	1.07	0.004	255.4	$400 \times 400 \times 8$	$400 \times 400 \times 32$	30	1.86	1.77	0.16	6.12	N
R4	12	322.0	0.56	1.07	0.002	220.7	$400 \times 400 \times 8$	$400 \times 400 \times 32$	30	1.2	4.0	0.16	5.16	N
R5	12	322.0	1.33	0.87	0.004	44.7	$400 \times 400 \times 8$	$400 \times 400 \times 32$	3.6	0.5	4.9	0.16	1.24	Y
SR	24	35.0	8.0	0.20	0.083	6.3	$400 \times 400 \times 8$	$400 \times 400 \times 32$	10.0	0.5	1.20	0.16	0.30	Y
Z	24	35.0	4.0	0.52	0.077	13.1	$30 \times 30 \times 8$	$32 \times 32 \times 32$	30	1.0	0.46	0.16	0.46	N
Q	48	35.0	8.0	0.35	0.12	8.6	$20 \times 20 \times 8$	$32 \times 32 \times 32$	30	1.0	0.34	0.16	0.34	N
M	24	8.0	8.0	0.24	0.71	3.2	$8 \times 8 \times 8$	$32 \times 32 \times 32$	30	1.0	0.25	0.16	0.25	N

<sup>(a)</sup> Values are derived using  $Q_g^* = 1$ .

<sup>(b)</sup> Values are derived using  $Q_d^* = 1$ .

<sup>(c)</sup> GI = Gravitational Instability. Y means  $\max \rho_d$  increases by orders of magnitude over a few dynamical times, and N means it does not.

<sup>(d)</sup> See Figure 11.

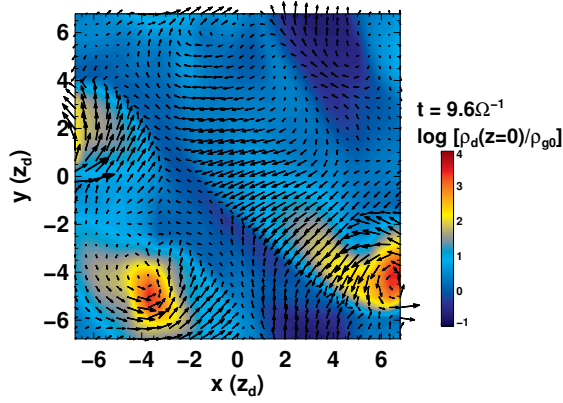


Fig. 7.— Snapshot of the midplane for run S = STD32 at  $t = 9.6\Omega^{-1}$ . The largest in-plane velocity shown is  $2.16 c_g$ .

illustrated in Figure 7 with a snapshot of the midplane slice of STD32 at  $t = 9.6\Omega^{-1}$ . The dust clumps are seen spinning about their centers of mass as a consequence of angular momentum conservation.

### 3.2.2. Resolution and box size

Table 4 lists the parameters of experiments designed to test our choices for resolution, box size,

and grid-cell aspect ratio.

Figure 8 shows how varying the resolution changes the evolution of our standard, gravitationally unstable run (STD32 — also labeled S in Table 3). We use again the simple metric of  $\max \rho_d$  vs.  $t$ . Broadly speaking, the runs STD16, STD32, STD64 are all “acceptable” insofar as they all yield increases in  $\max \rho_d$  by orders of magnitude within several dynamical times ( $t \lesssim 8\Omega^{-1}$ ). By contrast, the lowest resolution run, STD8, is unacceptable. Thus, the minimum acceptable resolution appears to be  $\sim 2$  cells per scale length  $z_d$  in the vertical direction (cf. Nelson 2006 who found that a minimum of four smoothing lengths per scale height is required for SPH simulations), and  $\sim 8$  cells per critical wavelength  $\lambda_c$  in the horizontal directions. Our standard choices for resolution — as well as the resolutions characterizing all our “science” runs, listed in Table 3 and discussed in §3.2.3 — satisfy these minimum requirements by a safety factor of 2.

Examining Figure 8 more critically, we see that the maximum value attained by  $\max \rho_d$  has not converged with resolution. Increasing the resolution enables us to resolve ever higher densities in the collapsing clumps. Another point of concern is the non-uniform aspect ratios of individual grid cells, which ranges from  $x:y:z \approx 2:2:1$  to  $4:4:1$  over

TABLE 4  
3D SIMULATION PARAMETERS TO TEST BOX SIZE AND RESOLUTION

Name	$L_x \times L_y \times L_z(z_d^3)$	Resolution	GI <sup>(a)</sup>	Duration ( $\Omega^{-1}$ )
STD32(S)	$14 \times 14 \times 8$	$32 \times 32 \times 32$	Y	9.8
STD8	$14 \times 14 \times 8$	$8 \times 8 \times 8$	N	30.0
STD16	$14 \times 14 \times 8$	$16 \times 16 \times 16$	Y	11.0
STD64	$14 \times 14 \times 8$	$64 \times 64 \times 64$	Y	11.0
U32	$14 \times 14 \times 8$	$56 \times 56 \times 32$	Y	10.0
LZ2	$14 \times 14 \times 2$	$32 \times 32 \times 8$	N	30.0
LZ4	$14 \times 14 \times 4$	$32 \times 32 \times 16$	Y	8.5
LZ6	$14 \times 14 \times 6$	$32 \times 32 \times 24$	Y	8.0
LZ10	$14 \times 14 \times 10$	$32 \times 32 \times 40$	Y	9.0
LZ14	$14 \times 14 \times 14$	$32 \times 32 \times 56$	Y	10.5
LXY6	$6 \times 6 \times 8$	$16 \times 16 \times 32$	N	30.0
LXY10	$10 \times 10 \times 8$	$24 \times 24 \times 32$	Y	8.6
LXY20	$20 \times 20 \times 8$	$48 \times 48 \times 32$	Y	8.7

<sup>(a)</sup> GI = Gravitational Instability. Y means  $\max \rho_d$  increases by orders of magnitude over a few dynamical times, and N means it does not.

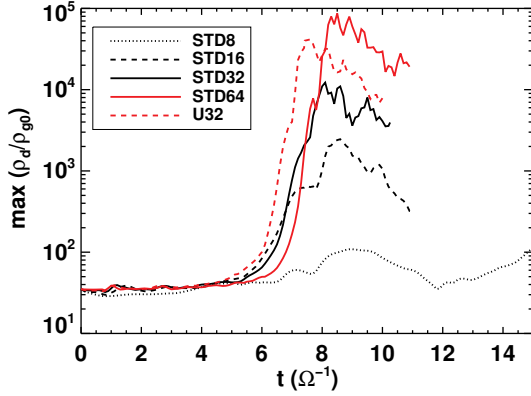


Fig. 8.— Time evolution of the maximum dust density for the resolution study (see Table 4).

our set of science simulations (Table 3). The run U32 is characterized by perfectly cubical grid cells (1:1:1); the evolution is similar to STD32, but is characterized by an earlier onset of gravitational instability, and stronger density fluctuations. This comparison suggests that our science runs with non-cubical grid cells are biased slightly against gravitational instability.

We next investigate how box size affects our results. For all box size experiments, the spatial resolution is kept at its standard value (32 grid cells per  $14z_d$  in either horizontal direction, and 4 grid

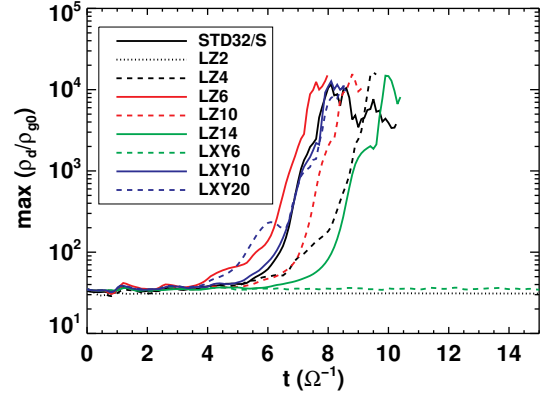


Fig. 9.— Time evolution of the maximum dust density for our box size tests (see Table 4).

cells per  $z_d$  in the vertical direction). Runs LZ2 through LZ14 vary box height  $L_z$  while keeping  $L_x$  and  $L_y$  fixed at their standard (STD32 = S) values. As Figure 9 reveals, box heights of 4–14 $z_d$  yield comparable results, while a box height of 2 $z_d$  is unacceptable. For the most part, increasing the box height seems to delay the onset of gravitational instability, with LZ4 being the exception to this rule.

Our 2D simulations indicated that  $L_x$  and  $L_y$  must be large enough to encompass at least one critical wavelength  $\lambda_c$ . Our 3D simulations bear



out this same requirement. Figure 9 shows that run LXY6, for which the box size is just under one critical wavelength, does not exhibit gravitational instability, unlike its bigger box counterparts.

To summarize our findings in this subsection: (1) The simulation box should be at least  $4z_d$  tall ( $2z_d$  above and below the midplane). (2) Each horizontal dimension must be longer than one critical wavelength  $\lambda_c$  as given by equation (30). (3) Simulations require a vertical resolution of  $\gtrsim 2$  grid cells per scale length  $z_d$ , and a horizontal resolution of  $\gtrsim 8$  grid cells per critical wavelength. (4) Individual grid cells that have increasingly non-uniform aspect ratios (squatter vertically than horizontally) tend to suppress gravitational instability, but the bias is minor and aspect ratios up to 4:4:1 appear acceptable. All of our science simulations (Table 3; §3.2.3) satisfy these requirements, in some cases by factors of 2.

### 3.2.3. Criteria for gravitational collapse

Table 3 lists the simulations designed to test which of the various proposed criteria for gravitational instability is the best predictor of collapse. Figures 2 and 10 describe the initial dust and gas profiles, while Figure 11 displays the results using our simple diagnostic of  $\max \rho_d$  vs. time.

First consider runs S and R1–R5, and ask whether these runs favor  $\rho_I^*$  or  $\rho_{II}^*$  for the density required for gravitational collapse. Because dust is a major component of our disks, we do not expect  $\rho_I^*$  — which is strictly valid only for pure gas disks — to be a good predictor. Indeed in all six of these runs, the midplane density  $\rho_0$  exceeds  $\rho_I^*$ , by factors of 7.5–30, yet only runs S and R5, and to a much lesser extent R1, exhibit collapse. All six runs indicate instead that  $\rho_{II}^*$  — equivalently,  $Q_d$  — is the better predictor, with the critical value

$$0.5 < Q_d^* < 0.9. \quad (33)$$

There is some concern that the comparison between runs R2–R5 and run S may not be fair because runs R2–R5 have a factor of  $\sim 2$  poorer spatial resolution in  $x$  and  $y$  compared to run S. This concern is allayed by run SR, which has the same physical parameters as S but is run with the box size and resolution of R3, and which turns out to behave qualitatively similarly to S (see Figure 11).

Our conclusion that  $\rho_{II}^*$  is relevant and that  $Q_d^*$  obeys (33) is supported further by runs Z, Q, and

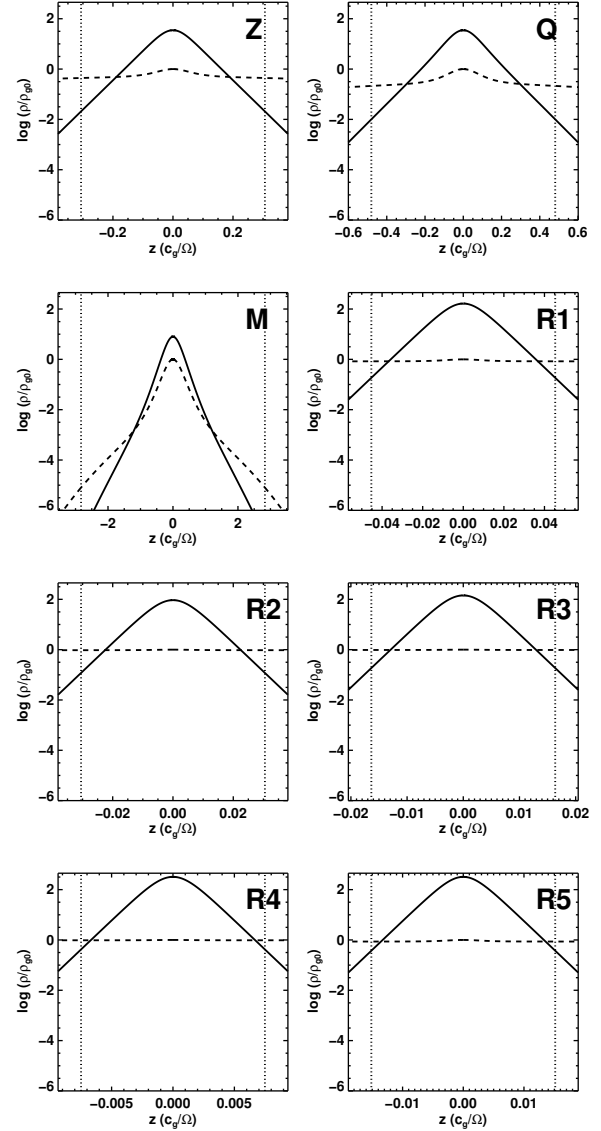


Fig. 10.— Initial conditions for our science runs which explore parameter space. Solid lines denote dust, and dashed lines denote gas. The vertical lines delimit the vertical boundaries of our simulation box.

M, each of which varies one of the three input parameters  $\Sigma_d/\Sigma_g$ ,  $Q_d$ , and  $\mu_0$ .

Although runs R2–R4 do not exhibit the dramatic growth in  $\rho_d$  shown by runs S, SR, and R5 — a result that we interpret to mean that  $\rho_{II}^*$  gives the correct criterion for gravitational collapse —

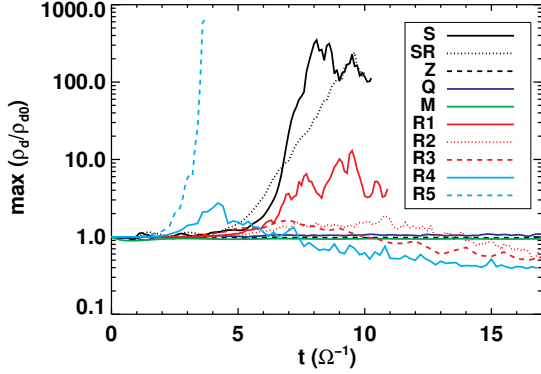


Fig. 11.— Time evolution of the maximum dust density in our science simulations. Only for runs S, SR, and R5 does  $\rho_0 > \rho_{\text{II}}^*$ , and indeed only those runs exhibit dramatic growth of the dust density due to gravitational instability.

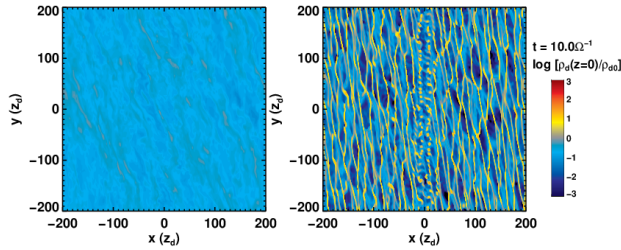


Fig. 12.— Snapshots of the dust density at the midplane for run R3 (left panel) and run SR (right panel). In R3, the initial midplane density  $\rho_0 > \rho_{\text{Sekiya}}^*$ , and there is some clumping, but it is much lower in amplitude compared to run SR, for which  $\rho_0 > \rho_{\text{II}}^*$ .

runs R2–R4 do show some clumping. Figure 12 compares snapshots of runs R3 and SR (performed with the same box size and resolution), taken at the same time  $t = 10\Omega^{-1}$ . Filaments do form in R3, although they are much weaker in density contrast compared to the filaments in SR. The mild growth shown in runs R2 and R3 might simply reflect the fact that their values for  $Q_d = 1.86$  are still too close to  $Q_d^*$  to suppress instability entirely. An alternative (and not mutually exclusive) possibility is that because  $\rho_0 > \rho_{\text{Sekiya}}^* = 0.60\rho^\dagger$  for runs R2–R4, the disk might be exhibiting the unstable (and formally incompressible) mode found by Sekiya (1983). Whatever the interpretation, the modest growth factors exhibited by R2–R4

seem unlikely to lead to planetesimal formation.

Finally, what about  $\rho_{\text{Roche}}^*$  vs.  $\rho_{\text{II}}^*$ ? Here runs R4 and R5 are the most telling. Both runs are characterized by the largest midplane densities  $\rho_0 > \rho_{\text{Roche}}^*$ , but only R5, for which  $\rho_0 > \rho_{\text{II}}^*$ , undergoes gravitational collapse (see Figure 11).

Table 5 summarizes how the various candidate critical densities relate to one another and to the midplane density for our science simulations. From Table 5,  $\rho_{\text{II}}^*$  emerges as the best predictor of collapse.

#### 4. SUMMARY AND DISCUSSION

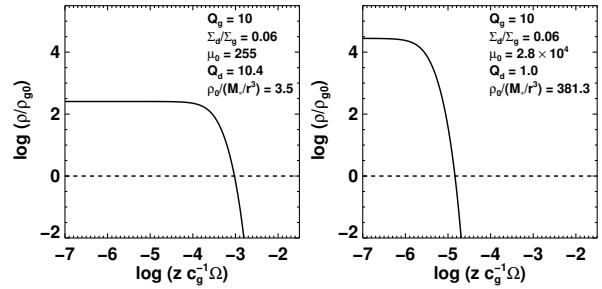


Fig. 13.— A tale of two particle sublayers, one of which is thinner and denser than the other. Dust density is plotted as a solid line, and gas density as a dashed line. The disks have identical masses and bulk metallicities, enhanced over those of the minimum-mass solar nebula by factors of 3–4. Left: Midplane density  $\rho_0 = \rho_{\text{Roche}}^* = 3.5\rho^\dagger$  and  $Q_d = 10.4$ . Right: Midplane density  $\rho_0 \approx \rho_{\text{II}}^* \approx 10^2 \rho_{\text{Roche}}^*$  and  $Q_d = 1$ . According to the results of our simulations, only the model in the right panel, having the thinner and denser sublayer, should be on the verge of gravitational collapse — in the limit that particles are aerodynamically perfectly coupled to gas. We argue in §4.1 that when the perfect coupling approximation breaks down, it may be possible for the disk on the left to undergo gravitational instability.

Dust grains settle toward the midplanes of protoplanetary disks, forming a sublayer of solid particles sandwiched from above and below by gas. Whether this sublayer can become thin enough and dense enough to undergo gravitational instability and fragment into planetesimals is an outstanding question. We have found in this work that the density threshold for gravitational collapse can be extraordinarily high — much higher even than the Roche density  $\rho_{\text{Roche}}^* = 3.5M_*/r^3$ ,

TABLE 5  
COMPARISON OF CRITICAL DENSITIES AND ACTUAL MIDPLANE DENSITY FOR SCIENCE SIMULATIONS

Name	Critical density relations	GI <sup>(a)</sup>
S	$\rho_I^* < \rho_{II}^* < \rho_{\text{Sekiya}}^* < \rho_0 < \rho_{\text{Roche}}^*$	Y
R1	$\rho_I^* < \rho_{\text{Sekiya}}^* < \rho_{II}^* < \rho_0 < \rho_{\text{Roche}}^*$	Y/N <sup>(b)</sup>
R2	$\rho_I^* < \rho_{\text{Sekiya}}^* < \rho_0 < \rho_{\text{Roche}}^* < \rho_{II}^*$	N
R3	$\rho_I^* < \rho_{\text{Sekiya}}^* < \rho_0 < \rho_{\text{Roche}}^* < \rho_{II}^*$	N
R4	$\rho_I^* < \rho_{\text{Sekiya}}^* < \rho_{\text{Roche}}^* < \rho_0 < \rho_{II}^*$	N
R5	$\rho_I^* < \rho_{\text{Sekiya}}^* < \rho_{II}^* < \rho_{\text{Roche}}^* < \rho_0$	Y

<sup>(a)</sup> GI = Gravitational Instability. Y means  $\max \rho_d$  increases by orders of magnitude over a few dynamical times, and N means it does not.

<sup>(b)</sup> See Figure 11.

where  $M_*$  is the mass of the central star and  $r$  is the orbital radius. To trigger collapse in the limit that dust particles are small enough to be tightly coupled to gas, the density  $\rho_0$  in the sublayer must be such that the Toomre stability parameter

$$Q_d \approx \left( \frac{\rho_{II}^*}{\rho_0} \right)^{1/2} \lesssim 1 \quad (34)$$

where

$$\rho_{II}^* \approx \frac{M_*}{2\pi r^3} \frac{Q_g}{Q_d^{*2}} \left( \frac{\Sigma_g}{\Sigma_d} \right)^2. \quad (35)$$

(For more precise relations, see equations 8, 13, and 33.) Here  $Q_g$  is the Toomre parameter for the ambient (and much thicker) gas disk,  $\Sigma_d/\Sigma_g$  is the ratio of surface densities of dust and gas (i.e., the height-integrated metallicity), and  $0.5 < Q_d^* < 0.9$  as measured from our simulations. For an astrophysically plausible disk having  $3\times$  the mass of the minimum-mass solar nebula ( $Q_g \approx 10$ ) and a bulk metallicity enriched over solar by a factor of 4 ( $\Sigma_d/\Sigma_g \approx 0.06$ ), the critical density

$$\rho_{II}^* \approx 1.3 \times 10^2 Q_d^{*-2} \rho_{\text{Roche}}^*. \quad (36)$$

Figure 13 portrays two sublayers — one for which  $\rho_0 = \rho_{\text{Roche}}^*$  and another, much thinner sublayer for which  $\rho_0 \approx \rho_{II}^* \approx 10^2 \rho_{\text{Roche}}^*$  ( $Q_d = 1$ ). The results of our simulations, performed in the limit of perfect aerodynamic coupling between particles and gas, indicate that only the latter, much denser disk is on the verge of fragmenting.

Qualitatively, such extraordinary densities are required for gravitational instability because gas

pressure renders the sublayer extremely stiff. Sound-crossing times for thin layers are easily shorter than free-fall times. We can examine the competition between stabilizing pressure, stabilizing rotation, and de-stabilizing self-gravity in both the horizontal (in-plane) and vertical directions. Horizontal stability is controlled by  $Q_d$ : when  $Q_d > Q_d^* \sim 1$ , all horizontal lengthscales  $\lambda \lesssim 2c_d^2/G\Sigma_d$  are stabilized by pressure, and all scales  $\lambda \gtrsim 2c_d^2/G\Sigma_d$  are stabilized by rotation, where  $c_d$  is the effective sound speed in the dust-gas mixture. At the same time, vertical stability is assured whenever the sound-crossing time across the vertical thickness of the sublayer  $2H_d$  is shorter than the free-fall time:

$$\frac{2H_d}{c_d} < \frac{1}{\sqrt{G\rho_d}} \quad (37)$$

which, after substituting  $H_d \approx \Sigma_d/2\rho_d$  and  $c_d \approx c_g \sqrt{\rho_g/\rho_d}$ , translates to

$$\left( \frac{\Sigma_d}{\Sigma_g} \right)^2 \frac{1}{Q_g} < \frac{\pi}{2} \quad (38)$$

which is easily satisfied for reasonable disk parameters.

Taken at face value, the higher density threshold  $\rho_{II}^*$  that our work establishes argues against using aerodynamically well-coupled particles to form planetesimals. The Kelvin-Helmholtz instability probably stops the midplane density from approaching  $\rho_{II}^*$ . For disks enriched in both mass and metallicity (by factors of several over the

minimum-mass solar nebula), midplane densities might just barely reach  $\sim M_*/r^3$ , but probably will not exceed it (e.g., Chiang 2008, Lee et al. 2010a,b).<sup>2</sup>

#### 4.1. Limitations of Our Study, and Directions for Future Research

We have worked in the limit that the stopping times  $t_{\text{stop}}$  of particles in gas are small compared to all other timescales. But in reality, finite particle sizes imply finite  $t_{\text{stop}}$  (see Figure 1). When the assumption of infinitesimal stopping time breaks down, new effects may appear that might lower the threshold for gravitational instability.

One such effect is as follows. Consider again the competition between stabilizing pressure and de-stabilizing self-gravity (in either the vertical or horizontal directions). A major reason why the sublayer so strongly resists collapse is that sound waves travel quickly across it. We have taken the sound speed for our dust-gas suspension to be  $c_d = c_g/\sqrt{1 + \rho_d/\rho_g} \approx c_g\sqrt{\rho_g/\rho_d}$  (equations 9 and 10). But this presumes that particles are perfectly coupled to gas. If the sound-crossing time across some scale  $\lambda$  were to become shorter than the particle stopping time, i.e., if

$$\frac{\lambda}{c_d} \approx \frac{\lambda}{c_g} \sqrt{\frac{\rho_d}{\rho_g}} < t_{\text{stop}} \quad (39)$$

then our use of  $c_d \approx c_g\sqrt{\rho_g/\rho_d}$  would be invalid. Particles on scales  $\lambda$  would lose support from gas pressure and become susceptible to gravitational instability.

To get a sense of where in parameter space this instability may lie, we normalize  $\lambda$  to the full vertical thickness of the sublayer:

$$\lambda \equiv 2H_d \hat{\lambda} = \frac{\hat{\lambda} \Sigma_d}{\rho_d}. \quad (40)$$

where  $\hat{\lambda}$  can take any value (larger than or smaller than unity). Then equation (39) for the loss of pressure support translates to a midplane density (dominated by dust) of

$$\rho_0 \approx \rho_d \gtrsim \frac{2 M_*}{\pi r^3} \left( \frac{\Sigma_d}{\Sigma_g} \right)^2 \frac{\hat{\lambda}^2}{Q_g (\Omega t_{\text{stop}})^2} \quad (41)$$

<sup>2</sup>See, however, Sekiya (1998), who found that self-gravitating, non-rotating sublayers having constant Richardson number could develop cusps of infinite density at the midplane.

where  $\Omega$  is the Kepler orbital frequency. For self-gravity to resist tidal disruption,  $\rho_d = \rho_{\text{Roche}}^* = 3.5 M_*/r^3$ . Substituting this requirement into (41), we find that

$$\begin{aligned} \Omega t_{\text{stop}} &\gtrsim \left( \frac{2}{3.5\pi} \right)^{1/2} \left( \frac{\Sigma_d}{\Sigma_g} \right) \frac{\hat{\lambda}}{Q_g^{1/2}} \\ &\gtrsim 8 \times 10^{-3} \left( \frac{\Sigma_d/\Sigma_g}{0.06} \right) \left( \frac{\hat{\lambda}}{1} \right) \left( \frac{10}{Q_g} \right)^{1/2} \end{aligned} \quad (42)$$

for particles on scales  $\hat{\lambda}$  to decouple from sound waves. For  $\hat{\lambda} = 1$ , requirement (42) could be fulfilled by particles having sizes of a few millimeters to a few centimeters at distances of 1–10 AU (Figure 1 — but note that the curves in the figure need to be adjusted by factors of a few for mass-enriched nebulae). For  $\hat{\lambda} < 1$ , even smaller particles could lose pressure support and collapse gravitationally.

Future simulations that include finite particle stopping times could try to find such an instability. A complication would be that accounting for finite  $t_{\text{stop}}$  would introduce the streaming instability, which could prevent the dust density from attaining the Roche value — see, e.g., runs R21-3D and R41-3D in Figure 5 of Bai & Stone (2010), for which  $\Omega t_{\text{stop}} \leq 0.1$  and  $\rho_d < \rho_{\text{Roche}}^*$ . To find the instability that we are envisioning, one would have to restrict  $\Omega t_{\text{stop}}$  to small enough values to suppress the streaming instability — thereby permitting the setting of grains into sublayers for which  $\rho_d = \rho_{\text{Roche}}^*$  — while at the same time keeping  $\Omega t_{\text{stop}}$  large enough to satisfy (42) and nullify pressure support.

We are grateful to Eve Ostriker and Andrew Youdin for discussions. Section 4.1 was inspired by discussions with Eve that clarified the limitations of our study and pointed the way to a possible new route to gravitational instability. We thank Xuening Bai, Chang-Goo Kim, Eve Ostriker, Ian Parrish, and Jim Stone for help in augmenting *Athena*. The simulations were performed with the Berkeley cluster Henyey, which was made possible by a National Science Foundation Major Research Instrumentation (NSF MRI) grant. Financial support for the authors was provided by the Berkeley Center for Integrative Planetary Science, the Berkeley Theoretical Astrophysics Center, and

grants from NSF (AST-0909210) and NASA Origins.

## REFERENCES

- Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., Rosenfeld, K. A., Öberg, K. I., Birnstiel, T., Espaillat, C., Cieza, L. A., Williams, J. P., Lin, S.-Y., & Ho, P. T. P. 2012, *ApJ*, 744, 162
- Bai, X.-N. & Stone, J. M. 2010, *ApJ*, 722, 1437
- Binney, J. & Tremaine, S. 2008, *Galactic Dynamics: Second Edition*, 2nd edn. (Princeton University Press)
- Birnstiel, T., Ricci, L., Trotta, F., Dullemond, C. P., Natta, A., Testi, L., Dominik, C., Henning, T., Ormel, C. W., & Zsom, A. 2010, *A&A*, 516, L14
- Chandrasekhar, S. 1987, *Ellipsoidal Figures of Equilibrium*, 1st edn. (Dover Publications, New York)
- Chiang, E. 2008, *ApJ*, 675, 1549
- Chiang, E. & Youdin, A. 2010, *Annual Reviews of Earth and Planetary Science*, 38
- Gammie, C. F. 2001, *ApJ*, 553, 174
- Goldreich, P. & Lynden-Bell, D. 1965, *MNRAS*, 130, 125
- Johansen, A., Oishi, J. S., Low, M., Klahr, H., Henning, T., & Youdin, A. 2007, *Nature*, 448, 1022
- Johnson, B. M. & Gammie, C. F. 2003, *ApJ*, 597, 131
- Kim, C.-G., Kim, W.-T., & Ostriker, E. C. 2011, *ApJ*, 743, 25
- Koyama, H. & Ostriker, E. C. 2009, *ApJ*, 693, 1316
- Lee, A. T., Chiang, E., Asay-Davis, X., & Baranco, J. 2010a, *ApJ*, 718, 1367
- . 2010b, *ApJ*, 725, 1938
- Mamatsashvili, G. R. & Rice, W. K. M. 2010, *MNRAS*, 406, 2050
- Nelson, A. F. 2006, *MNRAS*, 373, 1039
- Okuzumi, S., Tanaka, H., & Sakagami, M.-a. 2009, *ApJ*, 707, 1247
- Sekiya, M. 1983, *Progress of Theoretical Physics*, 69, 1116
- . 1998, *Icarus*, 133, 298
- Stone, J. M. & Gardiner, T. 2009, *New A*, 14, 139
- Stone, J. M. & Gardiner, T. A. 2010, *ApJS*, 189, 142
- Stone, J. M., Gardiner, T. A., Teuben, P., Hawley, J. F., & Simon, J. B. 2008, *ApJS*, 178, 137
- Toomre, A. 1964, *ApJ*, 139, 1217
- Toomre, A. 1981, in *Structure and Evolution of Normal Galaxies*, ed. S. M. Fall & D. Lynden-Bell, 111–136
- Toro, E. 1999, *Riemann Solvers and Numerical Methods for Fluid Dynamics* (Berlin: Springer)
- van Leer, B. 2006, *Comm. Comput. Phys.*, 1, 192
- Weidenschilling, S. J. 1977, *MNRAS*, 180, 57
- . 1980, *Icarus*, 44, 172
- . 2006, *Icarus*, 181, 572
- . 2010, *Meteoritics and Planetary Science*, 45, 276
- Youdin, A. N. & Goodman, J. 2005, *ApJ*, 620, 459
- Zsom, A., Ormel, C. W., Dullemond, C. P., & Henning, T. 2011, *A&A*, 534, A73